



Baltic Marine Environment Protection Commission

Working Group on the State of the Environment and Nature
Conservation

STATE & CONSERVATION
4-2016, 3MA-3

Schwerin, Germany, 11-15 April, 2016

Document title	Development of the COMBINE reporting format for phytoplankton
Code	3MA-3
Category	INF
Agenda Item	3MA– Environmental monitoring and data
Submission date	18.3.2016
Submitted by	ICES

Background

The reporting of phytoplankton data to the COMBINE database has been explored together with Lead Country representatives on HELCOM phytoplankton indicator development. The purpose has been to evaluate if the COMBINE database can support the HELCOM phytoplankton indicators and identify which fields of the reporting format are important/mandatory for HELCOM assessment purpose.

In relation to this the ICES Data Centre has prepared a phytoplankton view (draft version at <http://dome.ices.dk/views/phytoplankton.aspx>).

At a web meeting held 14 March 2016 it was concluded that it would be possible to use the COMBINE database for the phytoplankton indicators, however that it is important to get more reported data. In some cases only abundance has been reported. If only abundance is reported it is important that the PEG biovolume list is referenced, otherwise the biovolumes cannot be calculated. The conclusion will also be considered at the annual meeting of PEG (25-29 April 2016).

Action requested

The Meeting is invited to:

- to take note of the clarification on the use of the COMBINE database for HELCOM phytoplankton indicators currently in development.

Status of COMBINE data for phytoplankton indicators

This document aims at giving an overview of the COMBINE data in relation to the phytoplankton indicators

- Phytoplankton community composition as a food web indicator (candidate indicator)
- Seasonal succession of dominating phytoplankton groups (pre-core indicator)

Representatives of Lead Country Finland on the first mentioned indicator, have prepared a list of database issues connected with the phytoplankton indicator work (Appendix 1). In relation to this the ICES Data Centre have prepared a phytoplankton view (draft version at <http://dome.ices.dk/views/phytoplankton.aspx>) trying to address some of the issues in relation to harmonisation of data and the integration with the PEG biovolume list (http://www.ices.dk/marine-data/Documents/ENV/PEG_BVOL.zip). Additionally the phytoplankton view can be used to evaluate the extent that the COMBINE database can support the phytoplankton indicator(s). A full description of information included in the view is included in Appendix 2. Note that not all reported data to the COMBINE database contain values for all the view fields, therefore the view has a number of empty fields where submitted data cannot supply the information included in the view.

At a web meeting (14/3-2016), the database issues were discussed. The meeting participants included a representative from Lead Country Finland (Sirpa Lehtinen), on the first mentioned indicator, a representative from university of Tartu, Estonia that have been working on a harmonised phytoplankton database for the Baltic region (Kalle Olli), and ICES Data Centre. Below is a summary of the outcomes.

- It was generally agreed that [WoRMS](#) should be used for harmonisation of taxonomy. The PEG biovolume list is “almost” fully linked with WoRMS and this work should continue. Based on abundance number, biovolume and carbon biomass can be calculated using the PEG biovolume list. This requires that data (species names and size-class) are reported referencing the PEG biovolume list. At present only a little less than half of the phytoplankton data in the COMBINE database are reported referencing the PEG biovolume list. Biovolume and carbon biomass can alternatively be reported directly, but the preference is that species names and size-classes referencing the PEG biovolume list should be used when reporting data.
- It was noted that the newly drafted COMBINE database online phytoplankton view can solve the issues concerning missing information in the existing, general ICES data portal (<http://ecosystemdata.ices.dk/>) which was designed as an inventory. In the draft proposal view, data can be harmonised with regards to taxonomy in most cases and calculation of biovolume and carbon biomass can be made if data are reported referencing the PEG biovolume list. It is recommended that PEG looks into whether the calculation done using the PEG biovolume list has been done correctly, also to see if additional checks are needed to ensure data consistency.
- It was discussed how to facilitate phytoplankton reporting to the COMBINE database. ICES has developed a simple reporting format, under a paid request by OSPAR, as an Excel file. The Simplified Format is already in place for contaminant and biological effects data in biota, sediment and seawater. The plan is to extend this to community data during 2016. The Simplified Format was developed to meet the needs of those data submitters who have no possibility for using the ICES Environmental Reporting format (ERF3.2) or only have Excel-stored data. A first draft of a simplified reporting format for phytoplankton can be downloaded at http://ices.dk/marine-data/Documents/ENV/Environment_Formats.zip.
- It was noted that it would be useful to include information on salinity, temperature and chlorophyll a, as additional information for the phytoplankton community data. Most submitters do not provide those parameters with the phytoplankton data but they may have submitted them together with oceanographic data. It was agreed that ICES should look into how to include additional oceanographic data to be linked with the phytoplankton data. The linkage would likely have to be done by an algorithm that defines a ‘window’ of similar date, depth and geographic position.

- It was agreed that ICES should make a summary of the phytoplankton data in the COMBINE database per country to be presented at the STATE meeting. Data missing from the COMBINE database could possibly be identified by comparing the COMBINE database with the data compiled at the University of Tartu, Estonia. They are doing an update of their database by 15 April.

The meeting concluded that it would be possible to use the COMBINE database for the phytoplankton indicators, but that it is important to get more reported data. In some cases only abundance has been reported. If only abundance is reported it is important that the PEG biovolume list is referenced, otherwise the biovolumes cannot be calculated.

COMBINE database summary for phytoplankton

The COMBINE database currently (March 2016) contains 528682 measurements (8611 samples) for phytoplankton. Out of these 230812 measurements (2909 samples) are reported referencing the PEG biovolume list (~33%). For the indicator work, the time range for data needed is 2011 onwards. Below table summarises the current status of reporting per country from 2011 onwards. Total number for this time period amounts to 156979 measurements (2195 samples).

COMBINE phytoplankton data summary for 2011 onwards (March 2016)

Country	PARAM_desc	PEG_reference	minYear	maxYear	Samples	Measurements
Denmark	Abundance number (number counted)	No	2011	2014	721	18274
Estonia	Abundance number (number counted)	No	2014	2015	66	2199
Estonia	Biomass - carbon content	No	2014	2015	66	2191
Estonia	Biomass - wet weight	No	2014	2015	66	2199
Finland	Abundance number (number counted)	Yes	2011	2014	265	16389
Finland	Biomass - cell volume	Yes	2011	2014	265	16389
Germany	Abundance number (number counted)	No	2013	2014	95	2160
Germany	Biomass - carbon content	No	2013	2014	95	2128
Germany	Biomass - wet weight	No	2013	2014	95	2160
Lithuania	Abundance number (number counted)	Yes	2011	2014	95	7734
Lithuania	Biomass - wet weight	Yes	2011	2014	95	7734
Poland	Abundance number (number counted)	No	2011	2011	6	7
Poland	Abundance number (number counted)	Yes	2011	2014	231	15286
Poland	Biomass - carbon content	No	2011	2011	6	7
Poland	Biomass - carbon content	Yes	2011	2014	231	15273
Poland	Biomass - cell volume	No	2011	2011	6	7
Poland	Biomass - cell volume	Yes	2011	2014	231	15284
Sweden	Abundance number (number counted)	No	2011	2014	45	60
Sweden	Abundance number (number counted)	Yes	2011	2014	722	31498

Appendix 1 Phytoplankton database issues connected to indicator work

2.2.2016 Sirpa Lehtinen & Harri Kuosa (representatives of Lead Country Finland on the first mentioned indicator)

We foresee that the Phytoplankton community composition food web indicator could be fully operational to be used for the HOLAS II assessment by early 2017 provided that the phytoplankton long-term monitoring results from all HELCOM PEG countries are compiled and made available in a harmonized database by late 2016. But, the results of the Co-Leads need to be harmonized at least locally for testing this indicator candidate already during spring 2016.

We have been informed that it has been agreed by HELCOM State and Conservation to ensure that the COMBINE database can be used for the purpose of HELCOM assessments.

We are kindly suggesting that HELCOM might contact Timo Tamminen (timo.tamminen@environment.fi), Kalle Olli (kalle.oli@ut.ee) and Riina Klais (riina.klais@ut.ee) to discuss possible cooperation on phytoplankton database issues, since they have compiled a harmonized up-datable Baltic Sea wide phytoplankton database earlier as a project work.

Harmonized data – updating required annually

Running the long-term analyses requires that the results which are used together in the analyses are harmonized. Harmonization means that all taxon names, size-classes, biovolumes etc. fulfill the requirement of following one taxa and biovolume list. Since the HELCOM PEG taxa and biovolume list (which follows WoRMS) is updated annually, also the whole phytoplankton database should be up-dated annually according to the valid list. We are unaware if the ICES database is harmonized and up-dated, since checking the results from the ICES data portal to compare them with the results in our national database is not very easy. For quality checks, it would be also important to be able to follow the data processing.

Submitting results – not all data is presently available

At the moment, the process of reporting national results into the ICES database is time consuming since countries need to modify their results according to a certain template as extra work. In order to get PEG partners to report their results actively, the process of submitting results should be changed so that countries would be able to report their results preferably as they store them in their national databases, and the data harmonization would be done within the COMBINE database. This kind of solution has been developed by Kalle, Riina and Timo in their database.

Detailed supporting information required

It is very important that it is possible to get the data from the database in a very detailed form. Basic requirement is that all possible information on sampling, preservation, and microscopy is easily available for each sample, to be able to select only comparable results for the long-term analyses. Also other basic information such as the data provider, coordinates etc. should be easily available. Please, see the attached Excel file for an example of harmonized long-term results of one Finnish monitoring station.

Easy selection according to different taxonomical levels

It should be easy to pick the final results either on class, order, genus, species, or even size-class level, for example. Please, see the attached Excel file for an example of harmonized long-term results of one Finnish monitoring station.

Biomass results and abundance results

The biomass results (μg wet weight/liter) are the most important final results for this indicator. However, the HELCOM PEG countries use counting units and the abundance results are shown as counting units/liter

instead of cells/liter. As a counting unit may be either a cell, a certain number of cells or a certain length of a filament, the abundance results in cells/liter require converting within the database.

Appendix 2

ICES (DOME) Phytoplankton view **DRAFT**

Data reported for HELCOM COMBINE, OSPAR CEMP and other purposes

Description of fields

DOME output field	Description
tblSampleID	Reference to the sample (integer). Unique for a sample
Country	Country that has reported the measurement
MPROG	Monitoring programme (cf. MPROG). Multiple programmes are possible separated by "~" (ascii 126)
PURPM	Purpose of monitoring (cf. PURPM) . Multiple purposes possible separated by "~" (ascii 126)
STATN	Station name as reported
HELCOM_subbasin	Name of the HELCOM subbasin
RLABO	Reporting institute (cf. RLABO)
ALABO	Analytical laboratory (cf. RLABO)
SLABO	Sampling laboratory (cf)
HELCOM_L4_units	HELCOM assessment units (level 4) including WFD coastal and transitional waters
Latitude	Latitude of measurement (decimal degrees)
Longitude	Longitude of measurement (decimal degrees)
MNDEP	Minimum depth of sample (m)
MXDEP	Maximum depth of sample (m)
MYEAR	Monitoring year. Note that MYEAR may not always be identical to the year in date of sampling
DATE	Date of sampling (DD/MM/YYYY)
Year	Year of sampling (integer YYYY)
Month	Month of sampling (integer MM)
Day	Day of sampling (integer DD)
RLIST	Reference code list used for species (cf RLIST)
Species	Species/taxon name of examined specimen, as reported by submitter
SFLAG	Species flag (cf. SFLAG) . Multiple purposes possible separated by "~" (ascii 126)
SIZCL	Size class (value from reference in SIZRF)
SIZRF	Size class reference list (cf. SIZRF)
TRPHY	Trophic status of the species (cf. TRPHY)
STAGE	Developmental stage of the species (cf. STAGE)
COEFF	Coefficient
CPORT	Number of portions counted from split sample
NPORT	Number of portions in split sample
Value	Value measured
PARAM	Measured parameter (cf. PARAM)
PARAM_desc	Measured parameter full name
MUNIT	Measurement unit (cf. MUNIT)

final_value	For values that are reported as a number only (eg. 'nr' or 'g') the final concentration (eg. 'nr/l' or 'g/l') is calculated as number*COEFF. If value is reported as a concentration, final_value and value are identical
Final_value_unit	The unit of the final value
PEG_count_per_litre	The number of counting units per litre. This value is only calculated for abundance numbers reported with the size reference class list (SIZRF) 'PEG_BVOL'
PEG_volume_ug_per_litre	Biovolume (μg) per litre. This value is only calculated for abundance numbers reported with the size reference class list (SIZRF) 'PEG_BVOL'. The conversion from count per litre to biovolume per litre is done using conversion factors from the PEG biovolume list (cf PEG biovolume)
PEG_carbon_ug_per_litre	Carcon (μg) per litre. This value is only calculated for abundance numbers reported with the size reference class list (SIZRF) 'PEG_BVOL'. The conversion from count per litre to carbon (μg) per litre is done using conversion factors from the PEG biovolume list (cf PEG biovolume)
METOA	Method of analysis (cf. METOA)
MAGNI	Magnification
SMVOL	Total sample volume (l)
SDVOL	Sedimentation volume (ml)
METFP	Method of chemical fixation/preservation (cf. METFP)
QFLAG	Qualifier flag for measured value, i.e. "<" (cf. QFLAG). Multiple flags possible separated by "~" (ascii 126)
VFLAG	Validity flag, i.e. "S" suspect value (cf. VFLAG). Multiple flags possible separated by "~" (ascii 126)
FINFL	Factors potentially influencing guideline compliance and interpretation of data (cf. FINFL)
SMTYP	Sampler type (cf. SMTYP)
REFSK	Reference source or key (cf. REFSK)
FORML	Formula used in calculation (cf. FORML)
PEG_species	Species name as listed in the PEG biovolume list (cf PEG biovolume). The reported species name may be considered a synonym and can therefore be different
PEG_division	Taxonomic division as listed in the PEG biovolume list (cf PEG biovolume)
PEG_class	Taxonomic class as listed in the PEG biovolume list (cf PEG biovolume)
PEG_order	Taxonomic order as listed in the PEG biovolume list (cf PEG biovolume)
PEG_SFLAG	Species flag (cf. SFLAG) as listed in the PEG biovolume list (cf PEG biovolume)
PEG_STAGE	Developmental stage of the species (cf. STAGE) as listed in the PEG biovolume list (cf PEG biovolume)
PEG_AphiaID	Reference to a taxonomic unit in 'World Register of Marine Species' (cf WoRMS)
PEG_TRPHY	Trophic status of the species (cf. TRPHY) as listed in the PEG biovolume list (cf PEG biovolume)
PEG_geometric_shape	Geometric shape as listed in the PEG biovolume list (cf PEG biovolume)
PEG_formula	Formula for calculating biovolume as listed in the PEG biovolume list (cf PEG biovolume)

PEG_cells_no_per_counting_unit	The number of cells per counting unit as listed in the PEG biovolume list (cf PEG biovolume)
PEG_volume_um3_per_counting_unit	Biovolume (μm^3) per counting unit as listed in the PEG biovolume list (cf PEG biovolume)
PEG_carbon_pg_per_counting_unit	Carbon content (pg) per counting unit as listed in the PEG biovolume list (cf PEG biovolume)
PEG_sizerange	Size range of counting unit as listed in the PEG biovolume list (cf PEG biovolume)
tblUploadID	Reference to the uploaded submission (integer). Unique for a submission
tblSpotID	Reference to the sampling event (integer). Unique for a sampling event
tblParamID	Reference to the measurement (integer). Unique for the data table